# Valuing buildings energy efficiency through Hedonic Prices Method: are spatial effects relevant?

*Marta Bottero*, Marina Bravi**, Federico Dell'Anna***, Giulio Mondini*****

*key words:* buildings energy efficiency, green label, energy costs, Hedonic Prices Method (HPM), spatial econometric models, SAR, SEM

## Abstract

*The primary goal of this work is to employ a spatial econometric model joined with a basic Hedonic Prices Method (HPM) to estimate the implicit marginal price, as measure of willingness to pay for buildings energy performance in Turin City. The recent debate about environmental costs of energy waste justifies the implementation of different policies focused on buildings energy efficiency. The application of seven models on a large data-set of residential properties values shows the necessity to carefully control the coherence between spatial and econometric approaches. At the same time, findings of the exploration of an exemplary case study can help researchers and policy-makers in the definition of innovative urban models in the context of the post-carbon city.*

## 1. INTRODUCTION

In Italy, Regions and Municipalities are currently interested in encouraging buildings energy efficiency through targeted actions that however require a careful assessment. In this respect, the present work is devoted to estimating the social costs of buildings energy consumption with particular attention to spatial effects. The fact that energy consumption follows spatial patterns is quite obvious, since the building stock is not homogeneous, due to year of construction, structural and technological characteristics and, consequently, energy performance (Barthelmes *et al.*, 2016). The same possibilities of improving energy standards of historic buildings, rather than those built during the years of urban growth, are noticeably different. This does not mean that more recent buildings are necessarily high performing.

Until now, however, this problem has been considered more from a structural and technological point of view than from an economic one. So much that, in defining incentives policies, the structure of real estate ownership and market information are today probably underestimated. For instance, the introduction of the green label can be considered a true market signal for the buyer, a way to make transparent the information about buildings energy consumption and make the consumer more sensitive to the environmental issues. However, the need to mobilize the owners' willingness to pay to make the building stock more efficient appears evident today. Within this context, the present work is focused on the analysis of a case study, to say the least, exemplary from this point of view. The research, which is funded by the Interuniversity Department for Regional and Urban Studies and Planning of Politecnico di Torino, is also based on previous pilot experiments (Bottero & Bravi, 2014; Bottero *et al.*, 2016) and a collaboration with the Regional Authority and the Energy Center of Politecnico di Torino.

The work is organized into five sections. After a brief introduction, Section 2 provides a literature review of hedonic prices and spatial models, clarifying how the scholars have integrated the two approaches. The area under investigation and the methodology are described in Section 3 and 4. Finally, in Section 5, the results of the econometric application are discussed. Conclusions follow.

## 2. LITERATURE REVIEW

The Hedonic Pricing Method (HPM) is based on the idea that real estate properties are not homogenous goods (Rosen, 1974). Their market value is influenced by the presence of a bundle of attributes: locational, structural, temporal, geographical, and environmental. Each characteristic has an implicit price embodied in the selling price; the former is revealed only from observed values – revealed preferences – of differentiated products with a specific quantity of each attribute.

The method has a long experimental tradition and counts huge literature that cannot be summarized here. Freeman *et al.* (2014) noticed that economists have documented the relationship between housing prices and environmental amenities since before this link was recognized into the hedonic prices theory (Ridker, 1967). However, since it was established, the hedonic model was employed, under certain assumptions, to infer the marginal willingness to pay for properties attributes – first stage model – including environmental amenities. In light of this, a household maximizes its utility by simultaneously moving along each marginal price schedule, where this last can be interpreted as a household's willingness to pay for a unit of each attribute. In addition to marginal changes, HPM has been extended to value discrete changes in environmental amenities – second stage model – but this approach has not been widely practiced, due to a priori restrictions on household's preferences[1]. For example, an important assumption is that the urban area can be considered as a single market, where households must have perfect information on all alternatives and must be free to move into space. Obviously, this is an unrealistic assumption, because real estate markets are segmented and not transparent, and families do not have this possibility for several reasons, such as fixed costs, out loans and other subjective motivations.

From the formal point of view, the hedonic function $H$ is determined by different attributes, as represented in equation (1):

$$H = (S_i, N_i, Q_i, T_i) \tag{1}$$

where, for the property $i$, $S_i$ is a vector of structural attributes; $N_i$ is a vector of neighborhood attributes; $Q_i$ is a vector of environmental attributes; and $Ti$ is a vector of dummy variables reporting the sales period of time, as year, quarter or semester. Assuming now that the hedonic price function $H$ has been estimated for an urban area, its partial derivative with respect to any of its arguments, for example $Q$, gives the implicit marginal price as the additional amount that must be paid to move to a property with a higher quality level, *ceteris paribus*. If this function is nonlinear, the implicit marginal price of an attribute is not constant, but depends on its level and maybe – if interaction effects are considered – on the levels of other characteristics as well.

For empirical estimating, HPM relies on regression technique, which is criticized for a series of problems that can lead to biased estimates, such as functional form specification, spatial heterogeneity, spatial autocorrelation, housing quality change, multicollinearity, and heteroscedasticity. HPM has been continuously evolving, with the help of more powerful computation methods and evolutive techniques. One of these is the Geographic Information System (GIS), which allows, from the spatial point of view, more precise identification and valuation (Anselin, 1998). Another implementation is the evolution of the so-called big data, which let easy access, in a short while, to a huge mass of market information, reducing resources and time designated to data collection. The experiment presented here takes particular advantage of such a process[2].

The importance of space – or location – in determining real estate values is universally recognized. The introduction of spatial effects in HPM started from a reasoning about the autocorrelation of the error term in hedonic regression (Dubin, 1992). In this case, the neighborhood characteristics that cannot be captured by the analyst are considered responsible for causing biased estimates. Another issue is instead related to the adjacency effect, due to the nature of the real estate market (Can, 1992).

As a matter of fact, in a segmented and not perfectly competitive market, where information about prices and quantity is poor, buyers consult listing prices of nearby properties prior to making an offer. Similarly, sellers and agents use listing prices to determine a quotation and put

[1] When all households are similar with homogenous characteristics of income, the hedonic coefficient can be interpreted as marginal willingness to pay; but only in extreme cases, when all consumers have identical incomes and utility functions, the implicit marginal price curve is identical to the inverse demand function for an attribute.

[2] The Turin City database used in the present research is continuously implemented thanks to the collaboration with a large online real estate agency (www.immobiliare.it) and the possibility of accessing big data. A special thanks goes thus to immobiliare.it for the positive and continuous collaboration.

the good on the market. Particularly in boom phases, the sellers can drive the market, and buyers are more likely to accept quotations of similar properties recently sold. Overly optimistic buyers unwittingly reinforce spatial dependency of prices. Conversely, in bust phases, sellers may be hesitant to sell a property at a lower price than what they perceive as fair. This second kind of behavior can contribute to weakening the spatial effect, as well as reducing sales (Hyun & Milcheva, 2018). In other words, spatial economic phenomena can be explained much more by behavioral economics than by market equilibrium theory. Consistently, from the real estate appraisal point of view, the market value is determined by similar and recent sold properties and it is based on the comparative principle: another fact supporting the presence of adjacency effects. On one hand, it is about what is generally considered *iustum pretium* from the appraiser point of view. On the other hand, the presence of a reservation price and its role in determining the decision to sell or not is well documented in real estate literature (Haurin *et al.*, 2010) when sellers are driving the market. Finally, from the supply side, housing attributes exhibit a high degree of spatial correlation; properties near the city center are typically older, larger – at least in Italy – and without garages or other complementary features. On the contrary, suburban properties are generally newer, smaller, and, compared to energy efficiency, they are generally more performing.

Krause & Bitter (2012) found an increasing use, starting from the 2000s, of advanced spatial models in HPM literature as one of the leading trends in the real estate appraisal field. For example, Huang *et al. (*2017) examined the spatial distribution of residential properties prices in Shanghai using 12,732 valid observations. The analysis results were used to recommend the spatial pattern to government administration for formulating policies on land use and urban planning. A considerable amount of literature has been published on the effect of the green spaces on real estate prices. Du & Huang (2018) employed and compared three different spatial models to investigate the amenity value of urban wetland on houses prices in Hangzhou (China), and found positive and heterogeneous values for proximity.

Moreover, a large and growing body of literature has investigated the influence of undesired externalities on houses prices. Recently, a model proposed by Cordera *et al*. (2018) estimated the presence of spatial relationships between real estate values and the presence of an industrial area in the province of Taranto (Italy). A spatial model was conducted in Nantes (France) in order to verify the effects of air pollution and noise exposure on the houses prices (Boennec & Salladarré, 2017). Among others, two important contributions, for the present work, are that of Won Kim *et al*. (2003) and Chong *et al*. (2003), where the attention is focused on the joint application of spatial econometric models and environmental valuation.

## 3. STUDY AREA

As previously mentioned, the case study considered in the present research is related to the city of Turin. Turin's urban area was chosen for two reasons. First of all, Turin's air quality is very low compared to other European cities. The decline in air quality has been documented in recent reports about the main European cities (WHO, 2016; Legambiente, 2018), where Turin is found as one of the worst, with very high level of PM10 and other fine dust deriving from the heating sector for 49% on an annual basis and for 75% in the winter period (Arpa Piemonte, 2016). Second, the area presents a great level of energy consumption related to urban density (6,930.5 inh. per sq. km) and road traffic flows. Besides, the fundamental role of buildings in $CO_2$ emissions and energy consumption has been widely recognized (Klessmann *et al.*, 2011), recalling the European Union policies on this issue. Not less important for identifying the case–study has been the availability of a large data-set, continuously implemented, of real estate ads, with listed prices – or quotation – and many characteristics of interest, among which the energy consumption and the green label stand out.

From the geographical point of view, Turin, with a population of 884,733 inhabitants, is the capital of the Piedmont region besides the metropolitan area of the same name. The real estate market of Turin is one of the largest among Italian cities, but with the lowest houses prices in absolute terms: about 50% less than average prices in Milan and about 19% more than average prices in Palermo, which have the highest and lowest prices in the country respectively. According to the data coming from the *Agenzia delle Entrate* (OMI, 2017), against a recovery in the total number of sales, which started from 2014, after a long fall due to the global financial crisis, the average price continues getting down. This could be due to an excess of supply on demand. As proof of this: first of all, the overproduction of new buildings, caused by the urban transformations that, between 1995 and 2015, have remodeled five million square meters of industrial areas; secondly, the demographic decline of the urban area and the impoverishment of the population, especially the young and weak groups, now more than ever oriented to the rental market. Another important fact to consider are the characteristics of the existing building stock from the point of view of energy performance and maintenance status. Considering 1977 as a reference point in time – a crucial year for the construction industry because the first rules on the buildings energy efficiency came into force – the current real estate market shows a very high percentage of properties built before this date (83.48% of sales[3]), while only 8.36% is represented by buildings realized in the last 10 years. As a result, recent overproduction would seem to be partly absorbed by a market characterized by a stock far below energy consumption standards. The location of new buildings, with higher performances, is today point-like

[3] Our elaboration based on proprietary data.

rather than concentrated in some areas; or rather, it is located where the developer's profitability expectations are higher and where lands are available. On the contrary, this work moves from the need to monetize the benefits coming from incentive policies focused on the more lacking and intensive energy consume areas.

## 4. METHODOLOGY

To attain the primary goal, this study explicitly considers spatial effects in estimating the hedonic model for buildings energy efficiency. A sample consisted of 15,295 properties subject to sale for which a bid price – listing price – was published on the main Italian real estate portal, in a period between 2015 and the first quarter of 2018, was employed. It covers the full urban area of Turin City, as is visible on the map shown in Figure 1. First of all, a set of explanatory variables – where the certified energy consumption represents environmental characteristic[4] – have been selected. Some preliminary tests allowed identifying nine explanatory variables (Table 1), plus the dependent one as total listing price.

Before illustrating and commenting on the results, it is maybe useful to highlight some well-known problems that normally occur in this kind of applications. Generally speaking, the HPM relies on regression technique, which is criticized by some authors for a series of econometric problems that can lead to biased estimates, such as functional form specification, spatial adjacency, spatial autocorrelation, market segmentation and properties quality changes over time (Palmquist, 2005). For the purposes of this study, the discussion will focus on the first three issues.

First of all, estimating results are sensitive to the choice of functional form, as economic theory gives no clear guidelines on how to select it. However, the study of real estate markets has shown that, as other well-known economic phenomena, the selling price variation frequently shows a nonlinear relationship with the main explanatory variables. Furthermore, the search for individual/household willingness to pay requires non-constancy along the implicit price function. Typically, nonlinear hedonic price regression models are specified by applying a simple parametric model through logarithmic data transformation, often tested with a generalized Box-Cox quadratic model. However, some scholars (Cassel & Mendelsohn, 1985) have criticized this method because it does not always lead to consistent and interpretable estimates. Instead, a substantial difference concerns the estimation algorithm choice and implicit marginal prices

[4] A property more than 30–years–old should consume, on average, in a year, from 180 to 200 kWh per sqm. A huge requirement considering that a property with green label classified as "B" – a minimum standard for new buildings – can consume on average between 30 and 40 kWh per sqm per year. Certified energy consumption is, consequently, the energy consumption in kWh per sqm per year attributed to a property by green label certification. The technical name is global Energy Performance index (EPgl), also called IPE, an architectural parameter that indicates how much energy is consumed so that the building – or the real estate unit – reaches the comfort conditions for winter heating, production of domestic hot water, summer cooling and artificial lighting.

**Table 1** - *Variables list*

| DEPENDENT AND INDEPENDENT VARIABLES | | | | | |
|---|---|---|---|---|---|
| **Variables:** | **Measure** | **Min** | **Max** | **Mean** | **St. Dv.** |
| Surface (sqm) | Scale | 20 | 578 | 90.73 | 46.88 |
| Energy (kWh/year) | Scale | 126.36 | 219,190 | 16,846.36 | 11,326.25 |
| Green label (A=1; B=2; C=3; D=4; E=5; F=6; G=7) | Ordinal | 1 | 7 | 4.86 | 1.68 |
| EPI (kWh/sqm) | Scale | 3.5 | 975 | 188.87 | 82.32 |
| Floor | Scale | 0 | 15 | 2.88 | 2.14 |
| Elevator (1=There is; 0=There is not) | Nominal | 0 | 1 | 0.73 | 0.44 |
| Maintenance status (0 = Poor / To be restored; 1 = Good; 2 = Restored; 4 = New / Under construction) | Ordinal | 0 | 3 | 1.49 | 0.83 |
| Market segment (0 = Low; 1 = Medium; 2 = High; 3 = Very high) | Ordinal | 0 | 3 | 1.29 | 0.73 |
| Year_17 (1=2017; 0=Otherwise) | Nominal | 0 | 1 | 0.07 | 0.26 |
| Year_18 (1=2018; 0=Otherwise) | Nominal | 0 | 1 | 0.26 | 0.44 |
| Total listing price (dependent) | Scale | 90,000 | 36,000.00 | 186,672.84 | 17,506.11 |
| Price sqm | Scale | 412.5 | 8000 | 1,859.15 | 864.75 |

calculation. For example, Ordinary Least Square (OLS) involves using transformed data, while the Maximum Likelihood Estimator (MLE) allows using the original ones. The typical setting is, in the second case, as follows:

$$P_j = \beta_0 \varkappa_{j1}^{\beta 1} \varkappa_{j2}^{\beta 2} \dots \varkappa_{jn}^{\beta n} \varepsilon_j \tag{2}$$

where $P_j$ represents the price, $\beta_i$ are the estimation coefficients and $\varkappa_{ji}$ are the variables under examination and $\varepsilon_j$ the error term.

Using MLE, the log–log form, exponential in the coefficients – usually employed to estimate the Cobb-Douglas production function – allows obtaining the implicit marginal prices by calculating the following incremental ratio:

$$\delta P_j / \delta \varkappa_{ji} = (P_j / \varkappa_i) \beta_i \tag{3}$$

where $P_j$ is the estimated price using the parameters of the best fitted model and $\varkappa_{ji}$ is the quantity of the characteristic under investigation, as, for instance, energy consumption. In addition to marginal prices estimating, this model helps to take into account the complementarity – or interaction effect – between real estate attributes, which can be tested following this pattern: a) if the sum of the exponents of the regression equation is equal to 1, there is complementarity between the characteristics; b) if the same is greater than 1, there is an incremental complementarity; c) if it is less than 1, there is a decremental complementarity. This model is therefore useful in estimating implicit prices and willingness to pay.

Other important issues in HPM applications are, as mentioned above, spatial dependency, or adjacency effect, and spatial autocorrelation. When the errors are spatially correlated due to unobserved variables or measurement errors in characteristics related to the location, the model – otherwise defined Spatial Error Model (SEM) – has to be specified as follow:

$$\begin{gathered} P = \beta_0 + \varkappa_1 \beta_1 + \dots \varkappa_n \beta_n + \varepsilon \\ \varepsilon = \lambda W \varepsilon + u \end{gathered} \tag{4}$$

where $W$ is the spatial weighted matrix, $\lambda$ is the spatial error coefficient, and $u$ is an uncorrelated error term. As observed in the literature, the definition of $W$ is based on a series of non-neutral steps if referred to the estimation results (Seya *et al.*, 2013). Among others, the most popular approaches used to build spatial weight matrix are k-nearest neighbors, inverse cut-off distance and contiguity between polygons. The choice is also depending on data GIS structure: geographical coordinates – latitude and longitude – of points, polygons, raster or other significative geo–political entities. Moreover, $W$ is assumed to be exogenous for the purpose of identification, or parameter interpretation, and its diagonal elements are usually set to zero, in order to avoid predicting itself. Finally, it is also normalized by rows to prevent singularity (Anselin, 1988).

Another popular specification is the spatial autoregressive with a spatial weighted error or Spatial Autoregressive Model (SAR), where:

$$\begin{gathered} P = \beta_0 + \rho WP + x_1 \beta_1 + \dots x_n \beta_n + \varepsilon \\ \varepsilon = \lambda W \varepsilon + u \end{gathered} \tag{5}$$

The term $\rho WP$ corresponds to a weighted average price of neighboring observations and the parameters $\lambda$ and $\rho$ are commonly known as the autocorrelation coefficients. Summarizing, in the SEM, $\rho = 0$, and in the SAR, $\lambda = 0$, so that the errors, $\varepsilon$, are independent and identically distributed. The SAR model also implies that there exist direct spillover effects between the prices of neighboring properties (Le-Sage and Pace, 2009). The presence of $\rho$ and the matrix $W$ has, however, a significant effect on the marginal implicit prices calculation. In this case, following Won Kim *et al.* (2003), the formula became:

$$\delta P_j / \delta x_{ji} = \beta_i (I - \rho W_P)^{-1} \tag{6}$$

It can be interpreted as follows. The house price in the location $j$ is not only affected by a marginal change of one characteristic – for instance, energy consumption – of the property, but also by the marginal changes of the neighbors. The total impact of a change in energy consumption is the sum of direct and indirect impacts. In other words, there is a price adjustment between neighboring properties mainly due to the reasons set out in Section 2. This formula does not apply to the SEM, where the error correction makes up for the omission of variables related to externalities and local public goods and implicit marginal prices are constant, supposing the functional form is linear.

Due to simultaneity, SEM and SAR models cannot be estimated using OLS; therefore, MLE or instrumental variables methods are used instead.

## 5. EXPERIMENTAL RESULTS

The experimental findings are summarized in the Tables 2-8. Tables 2 and 3 highlight the results of the linear and log-log models computed via OLS estimator. Table 4 summarizes the estimates of the nonlinear (multiplicative exponential) model computed via MLE, while Tables 5-8 show the results of the SEM and SAR models, the only ones testing the spatial effects. As known, the OLS algorithm is based on simple and straightforward assumptions that can be summarized as follows: there is absence of multicollinearity of explicative variables; the error terms are assumed normal and independently distributed, with mean 0 and constant variance (heteroskedasticity). In general, OLS is rather robust, that is, small violations of the model hypotheses do not invalidate the inference or the conclusions. More important violations for at least one of the hypotheses can instead lead to severely misleading conclusions in parameters estimating. Usually, in the real estate market field, the main violation concerns the absence of error correlation, an issue to which the advanced regression models have tried to put right. The residual analysis easily highlights this issue also in this case. As already highlighted, it was necessary to understand if this was due mainly to spatial effects.

From Tables 2 and 3, it is easy verifying the correctness of signs, amounts and significance – fifth and sixth columns – of the single coefficients, in addition to the model goodness of fit. The nine previously identified explanatory variables pass the statistical significance test and show appropriate amounts and signs. The linear model is the only one having an immediate monetary quantitative meaning. In other words, from these results, it is possible to deduct the value of the implicit marginal prices immediately. A remark should be made about the negative sign of the FLOOR variable that should be carefully interpreted. Living on a high floor can be viewed as an advantage – if the building is very high with bright windows and big terraces – or a disadvantage, especially if there are no elevators. From a different point of view, more insolation mitigates the consumption of heating but increases the need for cooling in summer. So, as expected, the complementarity between real estate characteristics remains an unsolved issue in regression models. In fact, linear models cannot take into account the interaction effects between different variables, which explains why the exponential multiplicative model performs better (Table 4).

Other important interaction effects to consider for this study are represented by the relationship between GREEN LABEL and ENERGY and between SURFACE and ENERGY. As previously mentioned, this last is the result of the multiplication of the annual certified energy consumption – Global Energy Performance index (EPgl) – by the property surface. Statistics about multicollinearity – ninth and tenth columns – could help to detect if interaction effects are underestimated with the models using OLS. Although there is no particular threshold of the variance inflation factor (VIF) value that unequivocally determines the presence of multicollinearity, the variables with the highest levels can be identified.

Other significative variables are MARKET SEGMENT and MAINTENANCE STATUS. As mentioned in Section 3, in Turin, at the moment, the residential real estate market is characterized by a high percentage of old and energy consuming properties, at least offered on the market, albeit not sold yet. It is straightforward to understand the importance of the building maintenance status and market segment for demand orientation. Among other things, the former incontestably attests the presence of segmentation in the urban real estate market. Finally, YEAR_17 and YEAR_18 account for the temporal variation in prices; as explained before, even if the sales number is increasing, the listing price does not stop its declining.

**Table 2** - *Regression model results - Linear model (OLS)*

| REGRESSION RESULTS - LINEAR MODEL (OLS) - OBSERVATIONS NUMBER = 15,295 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Dependent Variable: Total listing price | Coefficients (β) | Std. Error | Std. Coeffi-cient | t | Sig. | 95.0% Confidence Interval | | Collinearity Statistics | |
| Independent variables | | | | | | Lower Bound | Upper ound | Tollerance | VIF |
| (Constant) | –134,096.486 | 4,268.013 | | –31.419 | 0.000 | –142,462.301 | –125,730.672 | | |
| Surface sqm. | 3,082.848 | 22.292 | 0.819 | 138.294 | 0.000 | 3,039.153 | 3,126.542 | 0.388 | 2.576 |
| Energy (kWh/year) | –0.579 | 0.092 | –0.037 | –6.303 | 0.000 | –0.758 | –0.399 | 0.392 | 2.549 |
| Green label | –4,471.308 | 515.204 | –0.043 | –8.679 | 0.000 | –5,481.170 | –3,461.446 | 0.563 | 1.777 |
| Floor | –2,639.573 | 313.860 | –0.032 | –8.410 | 0.000 | –3,254.776 | –2,024.369 | 0.935 | 1.070 |
| Elevator | 2,816.049 | 926.153 | 0.012 | 3.041 | 0.002 | 1,000.680 | 4,631.419 | 0.847 | 1.181 |
| Maintenance status | 20,141.998 | 921.113 | 0.095 | 21.867 | 0.000 | 18,336.508 | 21,947.489 | 0.724 | 1.381 |
| Market segment | 34,680.438 | 1,036.096 | 0.145 | 33.472 | 0.000 | 32,649.567 | 36,711.309 | 0.727 | 1.376 |
| Year_17 | –6,656.225 | 2,531.003 | –0.010 | –2.630 | 0.009 | –11,617.292 | –1,695.158 | 0.97 | 1.031 |
| Year_18 | –10,291.377 | 1,517.842 | –0.025 | –6.780 | 0.000 | –13,266.528 | –7,316.226 | 0.968 | 1.033 |
| Std. Error of the Estimate | 80,522.653 | R Square | 0.7920 | Adjusted Rd Square | | 0.7919 | Durbin-Watson Test | | 1.8537 |

**Table 3 -** *Regression model results - Nonlinear model (OLS)*

| REGRESSION RESULTS - LOG - MODEL (OLS) - OBSERVATIONS NUMBER = 15,295 | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Dependent Variable: Total listing price | Coefficients (β) | Std. Error | Std. Coefficients | t | Sig. | 95.0% Confidence Interval | | Collinearity Statistics | |
| Independent variables: | | | | | | Lower Bound | Upper Bound | Tolerance | VIF |
| (Constant) | 6.4420 | 0.0439 | | 146.7939 | 0.000 | 6.3560 | 6.281 | | |
| Surface sqm. | 1.3016 | 0.0086 | 0.7630 | 150.7554 | 0.000 | 1.2847 | 1.3185 | 0.515 | 1.943 |
| Energy (kWh/year) | –0.0653 | 0.0058 | –0.0566 | –11.2496 | 0.000 | –0.0767 | –0.0359 | 0.521 | 1.919 |
| Green Label | –0.1803 | 0.0108 | –0.0728 | –16.6374 | 0.000 | –0.2015 | –0.1590 | 0.689 | 1.451 |
| Floor | –0.0246 | 0.0045 | –0.0204 | –5.4373 | 0.000 | –0.0335 | –0.0157 | 0.941 | 1.062 |
| Elevator | 0.1379 | 0.0065 | 0.0840 | 21.3057 | 0.000 | 0.1252 | 0.1506 | 0.849 | 1.178 |
| Maintenance status | 0.2527 | 0.0076 | 0.1334 | 33.2556 | 0.000 | 0.2378 | 0.2676 | 0.82 | 1.220 |
| Market segment | 0.3733 | 0.0083 | 0.1871 | 45.1479 | 0.000 | 0.3571 | 0.3895 | 0.768 | 1.303 |
| Year_17 | –0.0349 | 0.0060 | –0.0215 | –5.8115 | 0.000 | –0.0467 | –0.0231 | 0.967 | 1.034 |
| Year_18 | –0.716 | 0.0062 | –0.0428 | –11.5937 | 0.000 | –0.0837 | –0.595 | 0.968 | 1.033 |
| Std. Error of the Estimate | 0.3278 | R Squared | 0.7985 | Adjusted R Squared | | 0.7983 | Durbin-Watson test | | 1.8962 |

**Table 4 -** *Regression model results - Multiplicative exponential model (MLE)*

| REGRESSION RESULTS - MULTIPLICATIVE EXP. MODEL (MLE) | | | | |
|---|---|---|---|---|
| Dependent Variable: Total listing price | Coefficients (β) | std. Error | 95.0% Confidence Interval | |
| Independent variables: | | | Lower Bound | Upper Bound |
| (Constant) | 412.2232 | 13.7226 | 385.3253 | 439.1210 |
| Surface sqm. | 1.2883 | 0.0065 | 1.2755 | 1.3011 |
| Energy (kWh/year) | –0.0222 | 0.0044 | –0.0308 | –0.0137 |
| Green Label | –0.1239 | 0.0063 | –0.1362 | –0.1116 |
| Floor | –0.0253 | 0.0042 | –0.0336 | –0.0170 |
| Elevator | 0.0941 | 0.0075 | 0.0794 | 0.1088 |
| Maintenance status | 0.1660 | 0.0079 | 0.1506 | 0.1815 |
| Market segment | 0.5287 | 0.0114 | 0.5064 | 0.5510 |
| Year_17 | –0.0369 | 0.0091 | –0.0548 | –0.0190 |
| Year_18 | –0.0494 | 0.0058 | –0.0609 | –0.0380 |
| R Squared | 1 – (Residual Sum of Squares) / (Corrected Sum of Squares) = 0.8226 | | | Obs. Number = 15,295 |

Before applying spatial models, a global spatial correlation analysis has been performed (Figure 1). For this purpose, the choice of the spatial weights matrix is mandatory. As mentioned in Section 4, the choice of the contiguity weights is a fundamental step of the spatial analysis. In this case, the choice fell on the creation of a contiguity matrix of Thiessen's polygons built around points – observations – identified by their geographical coordinates, latitude and longitude (Figure 2). Assessing whether two polygons are contiguous requires the use of explicit spatial data structures to deal with the location and arrangement of the polygons themselves. For this purpose, the study employs the freeware software GeoDa™ and its functionalities.

More specifically, Figure 1 shows the Moran's Index. The variables are standardized so that the units in the graph correspond to standard deviations. The four quadrants in the graph provide a classification of four types of spatial autocorrelation: high-high (upper right), low-low (lower left), for positive spatial autocorrelation; high-low (lower right) and low-high (upper left), for negative spatial autocorrelation. The slope of the regression line is Moran's Index, showed at the top of the graph (Anselin, 1996). The index shows a discrete level of spatial autocorrelation, albeit not too high, due probably also to the presence of some outliers.

The goodness of fit of the four spatial models (Tables 5 - 8) seems to give support to the hypothesis that a spatial pattern of real estate values – and, accordingly, of the WTP for energy consumption – does matter. It could be particularly true for the SEM, where the spatial effect is due to a lack of appropriate and complete identification of the explanatory variables at the micro-territorial level. The choice of this model also allowed abandoning the idea that the WTP for buildings energy consumption is influenced by the values of neighboring properties, a hypothesis not fully consistent with the theory of hedonic prices. Moreover, the necessity to obtain a not constant variation of the implicit marginal prices over the energy consumption function drove the final choice on a nonlinear error correction model without the effect of the FLOOR variable.

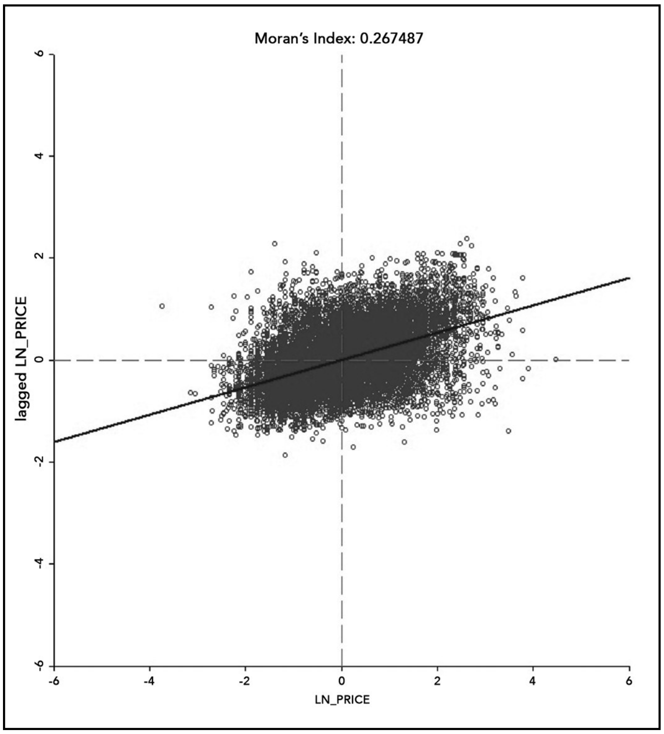


**Figure 1 -** *Moran's Index*

**Table 5 -** *Regression model results - Linear spatial error model (MLE)*

| LINEAR SPATIAL ERROR MODEL (MLE) | | | | |
|---|---|---|---|---|
| **Dependent Variable: Total listing price** / **Independent variables:** | **Coefficients (β)** | **Std. Error** | **z-values** | **Sig.** |
| Lag coefficient (λ) | 0.4463 | 0.0132 | 33.7930 | 0.0000 |
| Constant | –125.483,000 | 3,938.120 | –31.8635 | 0.0000 |
| Surface sqm. | 3,014.600 | 21.670 | 139.1150 | 0.0000 |
| Energy (kWh/year) | –0.494 | 0.088 | –5.6273 | 0.0000 |
| Green Label | –4,220.000 | 502,192 | –8.4032 | 0.0000 |
| Floor | –1,516.110 | 303,504 | –4.9954 | 0.0000 |
| Elevator | 5,841.510 | 1,551.220 | 3.7657 | 0.0002 |
| Maintenance status | 20,586.200 | 886,801 | 23.2140 | 0.0000 |
| Market segment | 29,194.300 | 1,004.810 | 29.0544 | 0.0000 |
| Year_17 | –7,572.070 | 2,421.290 | –3.1273 | 0.0018 |
| Year_18 | –9,182.330 | 1,455.660 | –6.3080 | 0.0000 |
| Std. Error of the Estimate | 76,855.2 | Akaike info criterion | | 387,898 |
| Log likelihood | –193,939.088 | Schwarz criterion | | 387,975 |
| R Squared | 0.810393 | Sigma squared | | 5.91E+09 |
| Spatial error dependence for weight matrix | DF | Value | Prob. | Obs. number = 15,295 |
| | 1 | 1,070.7764 | 0.0000 | |

**Table 6 -** *Regression model results - Linear spatial autoregressive model (MLE)*

| LINEAR SPATIAL AUTO-REGRESSIVE MODEL (MLE) | | | | |
|---|---|---|---|---|
| **Dependent Variable: Total listing price** / **Independent variables:** | **Coefficients (β)** | **Std. Error** | **z-values** | **Sig.** |
| Lag coefficient (ρ) | 0.2127 | 0.0069 | 30.7240 | 0.0000 |
| Constant | –160,314.000 | 3,939.340 | –40.6958 | 0.0000 |
| Surface sqm. | 2,970.150 | 21.923 | 135.4790 | 0.0000 |
| Energy (kWh/year) | –0,544 | 0.089 | –6.1364 | 0.0000 |
| Green Label | –3,826.050 | 498.595 | –7.6737 | 0.0000 |
| Floor | –1,931.870 | 303.823 | –6.3585 | 0.0000 |
| Elevator | 4,236.800 | 1,538.820 | 2.7533 | 0.0059 |
| Maintenance status | 20,356.100 | 890.93 | 22.8593 | 0.0000 |
| Market segment | 29,851.800 | 1,010.040 | 29.5551 | 0.0000 |
| Year_17 | –7,169.510 | 2,446.880 | –2.9301 | 0.0034 |
| Year_18 | –9,535.190 | 1,467.480 | –6.4977 | 0.0000 |
| Std. Error of the Estimate | 77,845.8 | Akaike info criterion | | 388,017 |
| Log likelihood | –193,997.000 | Schwarz criterion | | 388,101 |
| R Squared | 0.805473 | Sigma Squared | | 6.06E+09 |
| Spatial lag dependence for weight matrix | DF | Value | Prob. | Obs. number = 15,295 |
| | 1 | 954.1372 | 0.0000 | |

As previously mentioned, the implicit marginal price can be interpreted as the marginal WTP assuming the residential real estate market is in equilibrium. An important consideration to do is that the marginal benefits are capitalized into the property and they do not represent the annual revenue. As such, the marginal value is influenced by the length of time the owner expects to reside in the house, the price expected for this attribute when he will sell the property, the discount rate, and the future trend of energy costs. If energy costs are expected to rise in the future, the capitalized marginal benefits could fall. Conversely, capitalized marginal benefits could rise if energy costs are expected to fall.

**Table 7 -** *Regression model results - Nonlinear spatial error model (MLE)*

| NON LINEAR (LOG-LOG) SPATIAL MODEL (MLE) | | | | |
|---|---|---|---|---|
| **Dependent Variable: Total listing price** | **Coefficients (β)** | **Errore std.** | **z-values** | **Sig.** |
| **Independent variables:** | | | | |
| Lag coefficient (λ) | 0.5738 | 0.0113 | 50.9817 | 0.0000 |
| Constant | 6.5879 | 0.0410 | 160.5940 | 0.0000 |
| Surface sqm. | 1.2502 | 0.0080 | 155.8250 | 0.0000 |
| Energy (kWh/year) | –0.0552 | 0.0053 | –10.3462 | 0.0000 |
| Green Label | –0.1887 | 0.0107 | –17.5849 | 0.0000 |
| Floor | –0.0033 | 0.0041 | –0.7879 | 0.4307* |
| Elevator | 0.1340 | 0.0060 | 22.3165 | 0.0000 |
| Maintenance status | 0.2530 | 0.0069 | 36.7898 | 0.0000 |
| Market segment | 0.2988 | 0.0076 | 39.3463 | 0.0000 |
| Year_17 | –0.0277 | 0.0054 | –5.0965 | 0.0000 |
| Year_18 | –0.0577 | 0.0056 | –10.2731 | 0.0000 |
| Std. Error of the Estimate | 0.297667 | Akaike info criterion | | 6,974.92 |
| Log likelihood | –3,477.461892 | Schwarz criterion | | 7,051.28 |
| R Squared | 0.8337 | Sigma Squared | | 0.0886054 |
| Spatial error dependence for weight matrix | DF | Value | Prob. | Obs. number = 15,295 |
| | 1 | 2,324.6121 | 0.0000 | |

**Table 8 -** *Regression model results - Nonlinear spatial autoregressive model (MLE)*

| NON LINEAR (LOG-LOG) SPATIAL AUTO-REGRESSIVE MODEL (MLE) | | | | |
|---|---|---|---|---|
| **Dependent Variable: Total listing price**<br>**Independent variables:** | **Coefficients (β )** | **Std. Error** | **z-values** | **Sig.** |
| Lag coefficient (ρ) | 0.2899 | 0.0063 | 46.0262 | 0.0000 |
| Constant | 3.2782 | 0.0788 | 41.6220 | 0.0000 |
| Surface sqm. | 1.2337 | 0.0082 | 150.3770 | 0.0000 |
| Energy (kWh/year) | –0.0629 | 0.0054 | –11.6459 | 0.0000 |
| Green Label | –0.1477 | 0.0101 | –14.5868 | 0.0000 |
| Floor | –0.0103 | 0.0042 | –2.4461 | 0.0144 |
| Elevator | 0.1260 | 0.0060 | 20.9110 | 0.0000 |
| Maintenance status | 0.2505 | 0.0071 | 35.4373 | 0.0000 |
| Market segment | 0.3199 | 0.0077 | 41.3375 | 0.0000 |
| Year_17 | –0.0333 | 0.0056 | –5.9606 | 0.0000 |
| Year_18 | –0.0642 | 0.0057 | –11.1716 | 0.0000 |
| Std. Error of the Estimate | 0.304898 | Akaike info criterion | | 7,227.71 |
| Log likelihood | –3,602.86 | Schwarz criterion | | 7,311.7 |
| R Squared | 0.8226 | Sigma Squared | | 0.0929631 |
| Spatial lag dependence for weight matrix | DF | Value | Prob. | Obs. number = 15,295 |
| | 1 | 2,073.8239 | 0.0000 | |

Recalling that, in the case of buildings energy consumption, the unit of measurement is the annual kWh, the economic meaning is more immediate. It represents the annual cost for energy consumption that the owner is willing to pay for a house located in a certain urban area with specific structural characteristics, between which the green label is a relevant feature. As shown in Table 9, an annual WTP for energy consumption cost, for a house of 90.73 sqm, with a green label 'E', in 'good' state of maintenance, belonging to the 'medium' market segment, with elevator, listed today – property at the sample mean – is € 15,697.44. Obviously, this amount is greater than the actual annual costs an owner pays for heating, cooling and lighting his home. It represents the social cost to abandon the real estate stock in a bad average condition, rather than improve its quality and the energy performance. Considering that, in the better condition, this cost would be only € 117.74, the difference appears considerable and asks new interventions, among which the study of effective measures to encourage buildings – urban – energy efficiency is essential.

## 6. CONCLUSIONS

The need to develop incentive policies for improving buildings energy efficiency requires a careful assessment of the actual condition of existing building stock (D'Alpaos *et al.*, 2018; Bottero *et al.*, 2019). Current actions dedicated

**Table 9 -** *Implicit Marginal prices estimate*

| IMPLICIT MARGINAL PRICES FOR BUILDINGS ENERGY CONSUMPTION | | | | | |
|---|---|---|---|---|---|
| **Model without spatial effect** | **Mean** | **Std. Error** | **Model with spatial effect** | **Mean** | **Std. Error** |
| Linear (OLS) | –0.579 | –0.09178 | SEM (Lin) | –0.494 | –0.08780 |
| Log-log (OLS) | –0.9219 | –0.00570 | SAR (Lin) | –0.544 | –0.08873 |
| Multiplicative Exponential (MLE) | –0.9726 | –0.00205 | SEM (Log-Log) | –0.9318 | –0.00495 |
| Estimate for total building energy consumption (sample mean) = € 15,697.44 | | | SAR (Log-Log) | –0.9243 | –0.00561 |


**Figure 2 -** *Sample observations location*

to increasing the buildings energy efficiency include those that provide for an economic incentive, or a tax exemption, and those that introduce volumetric deductions or exemptions on the constraints of building regulations (Napoli *et al.*, 2017). In Italy, are currently in force: tax deductions of 65% for private parties, a thermal bill that provides incentives for both public and private subjects and volumetric deductions as required by national law. However, the need to involve privates in actions devoted to the requalification of the building stock calls for the need to develop and calibrate interventions at the urban scale (Becchio *et al.*, 2018). In this respect, the social costs and benefits estimating become crucial for justifying and giving priority to any intervention (Becchio *et al.*, 2019).

Moving from the need to assess the social cost of energy waste, this work attempted to estimate the differential in buildings energy performance in monetary terms. In spite of the complexity of an approach including spatial effects into the econometric model, the results are consistent and encouraging. At the same time, they signal the need to refine the analysis. With the aim of further developing this research to obtain increasingly reliable estimates, we can identify the following goals: a) in order to design interventions and incentives at the urban scale, the use of political-administrative units, such as cadastral or census zones, rather than single points, should provide more expendable and immediate findings; b) a more precise geographic representation of the estimates of the average cost per unit of energy consumption could help the Municipality to intervene by following a priority rank; c) time variable should be included in a more sophisticated fashion using an appropriate model on a bigger database able to take into account smaller time variations; c) among other things, a broader database might give the opportunity to refine the analysis on homogeneous market segments bypassing the assumption about a unique and in equilibrium real estate market. This last hypothesis reinforces the idea of continuing to work on *big data*, despite the partial loss of information that this entails. The quotation is not, in fact, equal to the selling price and it is necessary to take into account a certain percentage of overestimation of the economic effects. As is well known, this percentage varies with the market scenario; in times of boom, it is minimal – considering the short time of the property on the market – while, during busts, it is wider and leads to the choice to withdraw the property from the market. A further development of the research work would concern the estimate of this percentage in a random scenario.

Summarizing, this effort represents one of many steps of a broader research work devoted to reducing buildings energy consumption at the urban scale and, in this direction, to improving environmental quality.

* **Marta Bottero,** *Department of Regional and Urban Studies and Planning (DIST), Politecnico di Torino.*
e-mail: *marta.bottero@polito.it*
** **Marina Bravi,** *Department of Regional and Urban Studies and Planning (DIST), Politecnico di Torino.*
e-mail: *marina.bravi@polito.it*
*** **Federico Dell'Anna,** *Department of Regional and Urban Studies and Planning (DIST), Politecnico di Torino.*
*email: federico.dellanna@polito.it*
**** **Giulio Mondini,** *Department of Regional and Urban Studies and Planning (DIST), Politecnico di Torino.*
e-mail: *giulio.mondini@polito.it*

**References**

Anselin L. (1988). *Spatial econometrics: Methods and models*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Anselin L. (1996). *The Moran Scatterplot as an ESDA Tool to Assess Local Instability in Spatial Association*. In M. Fischer, H. Scholten, and D. Unwin (eds.), *Spatial Analytical Perspectives on GIS*. London, Taylor and Francis, pp. 111-125.

Anselin L. (1998). *GIS Research Infrastructure for Spatial Analysis of Real Estate Markets*, Journal of Housing Research, 9, 1, 113-156.

Arpa Piemonte (2016). Uno sguardo all'aria; available online: http://www.cittametropolitana.torino.it/cms/risorse/ambiente/dwd/qualita-aria/relazioni-annuali/relazione2016.pdf. Consultato maggio 2018.

Barthelmes V., Becchio C., Bottero M., Corgnati, S. (2016). *Cost-optimal analysis for the definition of energy design strategies: the case of a nearly-Zero Energy Building*. Valori e valutazioni 16, 61-76.

Becchio C., Bottero M., Corgnati S., Dell'Anna F. (2018). *Decision making for sustainable urban energy planning: an integrated evaluation framework of alternative solutions for a NZED (Net Zero-Energy District) in Turin*. Land Use Policy, 78, 803-817.

Becchio C., Bertoncini M., Boggio A., Bottero M., Corgnati S.P., Dell'anna F. (2019). *The Impact of Users' Lifestyle in Zero-Energy and Emission Buildings: An Application of Cost-Benefit Analysis*. Smart Innovation, Systems and Technologies, 100, pp. 123-131, DOI: 10.1007/978-3-319-92099-3_15.

Boennec R.L., Salladarré F. (2017). *The impact of air pollution and noise on the real estate market. The case of the 2013 European Green Capital: Nantes, France*. Ecological Economics, 138, 82-89.

Bottero M., Bravi M. Gasca E., Mondini G., Talarico A. (2016). *Building energy performance and real estate market value: an application of the Spatial Autoregressive (SAR) model*. In: Stanghellini S. Morano P. Bottero M. Oppio A.In: *Appraisal: From Theory to Practice. Results of SIEV 2015*. Green energy and technology, Springer, 221-230.

Bottero M., Bravi M. (2014). *Valuing benefits from building energy saving: an econometric approach, GEAM*. Geoingegneria Ambientale e Mineraria, 143, 15-24.

Bottero, M., D'Alpaos, C., Dell'Anna, F. (2019). *Boosting investments in buildings energy retrofit: The role of incentives,* Smart Innovation, Systems and Technologies, 101, pp. 593-600, DOI: 10.1007/978-3-319-92102-0_63.

Can A. (1992). *Specification and estimation of hedonic* housing price models, Regional Sciences and Urban Economics, 22, 453-474.

Cassel E., Mendelsohn R. (1985). *The Choice of Functional Forms for Hedonic Price Equations: Comment*. Journal of Urban Economics, 18, 2, 135-142.

Chong W.K., Tim T.P., Anselin L. (2003). *Measuring the benefits of air quality improvement: a spatial hedonic approach*. Journal of Environmental Economics and Management, 45, 24-39.

Cordera R., Chiarazzo E., Ottomanelli M., Dell'Olio L., Ibeas A. (2018). *The impact of undesirable externalities on residential property values: spatial regressive models and an empirical study*. Transport Policy. In press.

D'Alpaos C. Bragolusi P. (2018). *Buildings energy retrofit valuation approaches: State of the art and future perspectives*. Valori e Valutazioni, 20, 79-92.

Du X., Huang Z. (2018). *Spatial and temporal effects of urban wetlands on housing prices: Evidence from Hangzhou, China*. Land Use Policy, *73*, 290-298.

Dubin R. (1992). *Spatial Autocorrelation and Neighborhood Quality,* Regional Sciences and Urban Economics, 22, 433-452.

Freeman A.M. III, Herriges J.A., Kling C.L. (2003). *The Measurement of Environmental and Resource Values:*

*Theory and Methods*. Resources for the Future, Routledge, Washington DC.

Haurin D.R., Haurin J.L., Nadauld T., Sanders A. (2010). *List prices, sale prices and marketing time: an application to housing markets*. Real Estate Economics, 38, 659-685.

Huang Z., Chen R., Xu D., Zhou W. (2017). *Spatial and hedonic analysis of housing prices in Shanghai*. Habitat International, 67, 69-78.

Hyun D., Milcheva S. (2018). *Spatial dependence in apartment transaction prices during boom and bust*. Regional Science and Urban Economics, 68, 36-45.

Klessmann C., Held A., Rathmann M., Ragwitz M. (2011). *Status and perspectives of renewable energy policy and deployment in the European Union – What is need to reach the 2020 targets?* Energy Policy, 39, 7637-7657.

Krause A., Bitter C. (2012). *Spatial Econometrics, Land Values and Sustainability: Trends in Real Estate Valuation Research*. Current Research on Cities, 29, 19-25.

Legambiente (2018). *Mal'aria 2018. L'Europa chiama, l'Italia risponde?* See https://www.legambiente.it/.

LeSage J.P., Pace R.K. (2009) *Introduction to Spatial Econometrics*. Boca Raton, FL: CRC Press, Taylor and Francis Group.

Napoli G., Gabrielli L., Barbaro, S. (2017). *The efficiency of the incentives for the public buildings' energy retrofit. The case of the Italian Regions of the "Objective Convergence"*. Valori e Valutazioni, 18, 25-40.

OMI - Agenzia delle Entrate (2017). *Rapporto immobiliare 2017*. Il settore residenziale; available online: Nsilib/Nsi/Schede/FabbricatiTerreni/omi/ Pubblicazioni. http://www.agenziaentrate.gov.it/wps/content/.

Palmquist R.B. (2005). *Property Value Model*. In: K. Mäler & J.R Vincent (Eds.). *Handbook of environmental economics*, Vol. II, Amsterdam: North-Holland, 763-819.

Ridker R.G. (1967). *Economic Costs of Air Pollution: Studies in Measurement*. New York, Praeger.

Rosen S. (1974). *Hedonic Prices and Implicit Markets: Product Differentiation in Perfect Competition*, Journal of Political Economy, 82, 1, 34-55.

Seya H., Yamagata Y., Tsutsumi M. (2013). *Automatic Selection of a Spatial Weight Matrix in Spatial Econometrics: Application to a Spatial Hedonic Approach*. 43, 429-444 Regional Science and Urban Economics.

WHO (World Health Organization) (2016) WHO Global Urban Ambient Air Pollution Database; see http://www.who.int/phe/health_topics/outdoorair/data bases/cities/en/

Won Kim C.W., Phipps T.T., Anselin L. (2003). *Measuring the benefits of air quality improvement: a spatial hedonic approach*. Journal of Environmental Economics and Management, 45, 24.39.